

Trade-off between Feedback Load for the Channel State Information and System Performance in MIMO Communications

Daniel Sacristán-Murga¹ and Antonio Pascual-Iserte^{1,2} *

¹Centre Tecnològic de Telecomunicacions de Catalunya (CTTC)

²Department of Signal Theory and Communications, Universitat Politècnica de Catalunya (UPC)

email: daniel.sacristan@cttc.es, antonio.pascual@upc.edu

Abstract. The performance of multiple-input multiple-output (MIMO) communication systems is greatly increased by having channel state information (CSI) at the transmitter. In systems with no channel reciprocity, a limited feedback link is used to send the CSI from the receiver to the transmitter. However, the resources for the feedback link come at the expense of resources from the communications link. This paper studies the trade-off between accurate feedback and system performance for systems using different feedback techniques. The optimum feedback load is computed for different transmission schemes including time-division duplex (TDD) and frequency-division duplexing (FDD).

Key words: MIMO systems, feedback communication, quantization, limited feedback, multiuser systems.

1 Introduction

The use of multiple antennas at both the transmitter and the receiver is known to increase the performance of a communication system greatly, in terms of capacity [1] and resilience to fading [2]. In order to maximize this performance increase, complete knowledge of the propagation channel is required. This is easily obtained at the receiver by means of a training signal and channel estimation techniques. However, at the transmitter, this is not possible if channel reciprocity does not hold. In this scenario a dedicated feedback link can be used to send the required channel state information (CSI) from the receiver to the transmitter.

There has been extensive research on quantization and feedback techniques for transmitting the CSI from the receiver to the transmitter. However, in most

* The research leading to these results has received funding from the European Community's Seventh Framework Programme (FP7/2007-2013) under grant agreement no. 248267 (BuNGee) and no. 216715 (NEWCOM++). It has received further support from the Catalan Government under grant 2009 SGR 891, and by the Spanish Government under project TEC2008-06327-C03 (MULTI-ADAPTIVE).

cases the performance analysis is evaluated without taking into account the cost of using feedback. If we take this cost into account explicitly it turns out that, while using a large amount of feedback improves the quality of the CSI available at the transmitter, it is not optimum from a perspective of system performance since the remaining radio resources available for the data link are lower. This is because the differential gain obtained by each additional feedback bit is a decreasing function and, eventually, it gets lower than the cost of dedicating an additional bit to feedback. In this paper we show that using low feedback rate is better than not using feedback at all and also better than using large amounts of feedback.

The paper is organized as follows. The system and signal models are given in section 2. Section 3 presents a review of available feedback strategies, and then the performance evaluation of the different techniques follows in section 4. Section 5 is devoted to the study of the trade-off between performance and allocation of radio resources between the data link and the feedback link. Finally, section 6 concludes the paper.

2 System and signal models

Let us consider a flat fading MIMO channel with n_T and n_R transmit and receive antennas, respectively, represented at time instant n by $\mathbf{H}(n) \in \mathbb{C}^{n_R \times n_T}$. The n_R received signals at the same time instant, assuming a linear transmitter, can be expressed as

$$\mathbf{y}(n) = \mathbf{H}(n)\mathbf{B} \left(\hat{\mathbf{R}}_H(n) \right) \mathbf{x}(n) + \mathbf{w}(n) \in \mathbb{C}^{n_R}, \quad (1)$$

where $\mathbf{x}(n) \in \mathbb{C}^{n_S}$ represents the n_S streams of signals to be transmitted with $\mathbb{E} \{ \mathbf{x}(n)\mathbf{x}^H(n) \} = \mathbf{I}$, and $\mathbf{B} \in \mathbb{C}^{n_T \times n_S}$ is the linear transmitter matrix that must satisfy the mean transmit power constraint $\|\mathbf{B}\|_F^2 \leq P_T$ ($\|\cdot\|_F$ stands for the Frobenius norm). Note that we explicitly indicate that the transmitter depends on the available estimate of the channel Gram matrix $\hat{\mathbf{R}}_H(n)$, where the exact matrix is $\mathbf{R}_H(n) = \mathbf{H}^H(n)\mathbf{H}(n)$ (the optimum precoding matrix \mathbf{B} depends only on $\mathbf{R}_H(n)$ as proved in [3], [4]). The additive white Gaussian noise (AWGN) at the receiver is represented by $\mathbf{w}(n) \in \mathbb{C}^{n_R}$ with $\mathbb{E} \{ \mathbf{w}(n)\mathbf{w}^H(n) \} = \sigma_w^2 \mathbf{I}$.

As a performance criterion we consider the packet transmission rate and also the mutual information using a Gaussian code, which is expressed as $\log_2 \det \left(\mathbf{I} + \frac{1}{\sigma_w^2} \mathbf{B} \mathbf{B}^H \mathbf{R}_H \right)$, as will be explained later in section 5.

3 Review of feedback strategies

Since the performance of the transmission scheme with partial CSIT depends on the feedback strategy followed, we will briefly introduce the feedback techniques evaluated in this paper. We will consider both non-differential and differential types of feedback [5], [6].

3.1 Non-differential feedback

The non-differential algorithms are based on the quantization of the CSI using a codebook, which is a set of precoding matrices that is known to the transmitter and the receiver. The receiver evaluates the performance for all the elements (codewords) in the set, and sends to the transmitter the index of the codeword that provides the best performance given the current channel. The main advantage of this technique is its simplicity, while its main drawback is that it does not exploit the correlation in time present in most real channels. There are different techniques to generate the codebook, such as the ones based on packaging in the Grassmannian manifold [7] or iterative procedures like in [5]. The techniques described in these references perform a quantization of the subspaces that correspond to the right eigenvectors of the channel matrix and do not consider the information of the individual eigenvalues. Consequently, the power allocation between spatial modes is constrained to be uniformly distributed.

3.2 Differential feedback

A differential feedback strategy is based on a quantization of the difference between the last known value of CSI and the current value. Only the information related to the update in the CSI is fed back to the transmitter, and this improves greatly the performance in time correlated channels at the cost of some slight complexity increase. In this paper we will evaluate the differential algorithm explained in detail in [6]. This algorithm performs a quantization of the channel Gram matrix ($\mathbf{R}_H = \mathbf{H}^H \mathbf{H}$), which is an Hermitian positive definite matrix by construction. The set of Hermitian positive definite matrices has the geometry of a convex cone [8], and this technique exploits such geometry. The algorithm is composed of 3 steps, which are performed at every feedback update. As depicted in Fig. 1, the steps for feedback interval n are defined as:

- **Initial situation:** The receiver has a perfect knowledge of the current channel matrix $\mathbf{H}(n)$. Both the transmitter and the receiver know which is the last quantization of the channel Gram matrix sent through the feedback channel $\hat{\mathbf{R}}_H(n-1)$. At the first feedback transmission the algorithm starts from the cone vertex: $\hat{\mathbf{R}}_H(0) = \mathbf{I}$ (Fig. 1.a).
- **Step 1:** Both the receiver and the transmitter generate a common set of Q geodesic curves² $\mathbf{\Gamma}_i(t)$ ($i = 1 \dots Q$) on the cone, having all of them the same initial point and with orthogonal directions³ $\dot{\mathbf{\Gamma}}_i(t)|_{t=0}$, which are determined

² A geodesic curve is defined as the path connecting two points in the set with minimum distance and with all its points belonging to the set. Consequently, its expression depends on the geometry of the set.

³ The maximum number of orthogonal directions is given by the dimension of the set of Hermitian matrices, i.e., $Q \leq n_T^2$. If the number of feedback bits is higher than $\log_2(2n_T^2)$, the additional directions will have the orthogonality constraint relaxed.

by the \mathbf{C}_i matrices⁴ (Fig. 1.b,c).

$$\mathbf{\Gamma}_i(t) = \hat{\mathbf{R}}_H^{1/2}(n-1) \exp(t\mathbf{C}_i) \hat{\mathbf{R}}_H^{1/2}(n-1). \quad (2)$$

- **Step 2:** Each of these geodesic curves $\mathbf{\Gamma}_i(t)$ is used to generate two candidates ($\mathbf{\Gamma}_i(1)$ and $\mathbf{\Gamma}_i(-1)$) for the feedback in the next iteration (Fig. 1.d).
- **Step 3:** The receiver evaluates the cost function for each of the candidates, selects the candidate that minimizes the cost function and sends the corresponding index i_{FB} through the feedback channel to the transmitter (Fig. 1.e). The selected matrix will be used for the transmitter design and as the starting point in the next feedback computation (Fig. 1.f).

4 Effect of feedback on the performance

This section will characterize, from a simulations perspective, the performance improvement achieved using the differential and non-differential feedback techniques described in section 3.

We consider a random MIMO channel following a first-order autoregressive time variation model, which is described by the expression:

$$\mathbf{H}(n) = \rho \mathbf{H}(n-1) + \sqrt{1 - \rho^2} \mathbf{N}(n), \quad (3)$$

where matrices $\mathbf{H}(0)$ and $\mathbf{N}(n)$ are assumed to be independent and composed of i.i.d. zero-mean complex Gaussian entries with unit variance. The time correlation factor ρ models the variability of the channel and depends on the Doppler frequency f_D caused by the movement of the transmitter/receiver through the expression $\rho = J_0(2\pi f_D \tau)$ [9], where J_0 is the zeroth-order Bessel function of the first kind and τ corresponds to the time difference between consecutive feedback instants. Note that the case of an invariant channel corresponds to $\rho = 1$. The time correlation factor is usually expressed in the literature in terms of the normalized Doppler frequency $f_D \tau$, or f_D/f_{FB} , where f_{FB} is the frequency of feedback messages. The values for this parameter usually considered in the literature are $0.004 < f_D/f_{FB} < 0.01$ (see references [10, 11, 12]), which correspond to $0.999 < \rho < 1$.

Fig. 2 shows the performance in terms of mutual information in a system with $n_T = 3$, $n_R = 3$, a normalized Doppler frequency of $f_D/f_{FB} = 0.05$ and a transmitted power of 0.5W. The simulations are averaged over 1000 channel realizations and show an improvement in the mutual information of 20% using just one bit of feedback, and up to more than 50% in the case of 6 bits of feedback. The incremental gain obtained by the use of each additional feedback bit is plotted in Fig. 3. It is a decreasing function, which means that the first bit introduces a great gain and each successive bit used provides a smaller gain. In

⁴ The quantization step is related to the norm of \mathbf{C}_i and can be optimized for any scenario and mobility conditions.

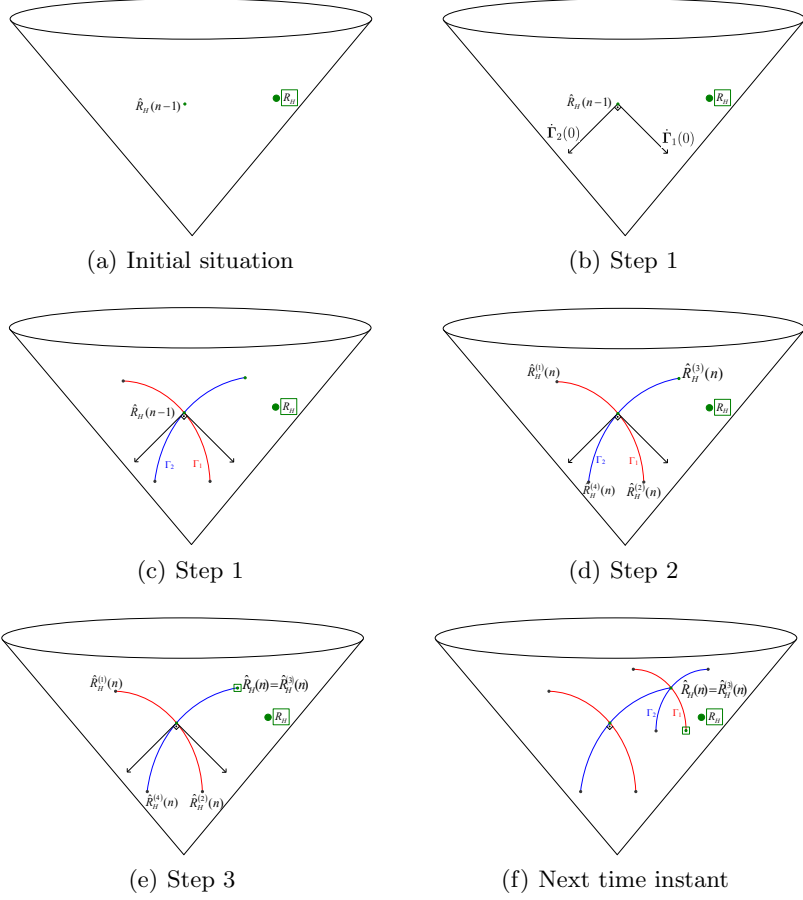


Fig. 1. Example of one feedback computation in a 2-bit differential quantization, using as optimization criterion the minimization of the geodesic distance to the actual channel Gram matrix \mathbf{R}_H .

these figures the differential feedback algorithm from [6] is compared to a differential quantization algorithm applied directly to the channel response matrix \mathbf{H} of the system instead of \mathbf{R}_H (so, the geometry of the convex cone corresponding to \mathbf{R}_H is not exploited).

The performance in terms of SNR is shown in Fig. 4, considering the following setup: $n_T = 3$, $n_R = 3$, $f_D/f_{FB} = 0.03$ and a transmitted power of 1W. The simulations are averaged over 1000 channel realizations and show a performance that increases with the number of bits of feedback. The differential strategy provides a better result than the non-differential one because it exploits the temporal correlation of the channel.

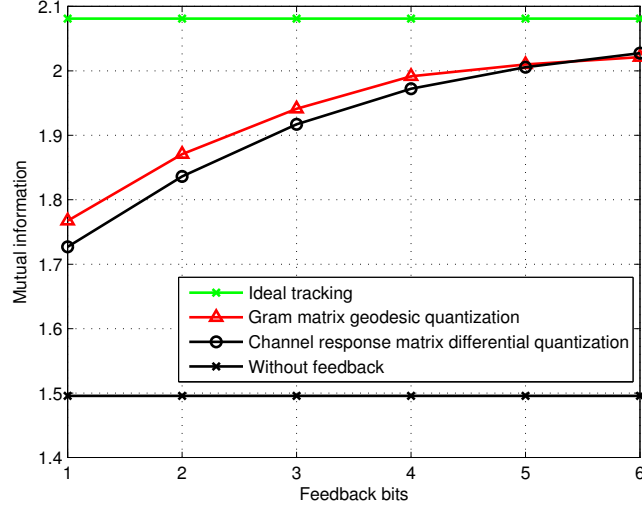


Fig. 2. Mutual information in a 3×3 channel with $P_{tx} = 0.5$, $f_D/f_{FB} = 0.05$.

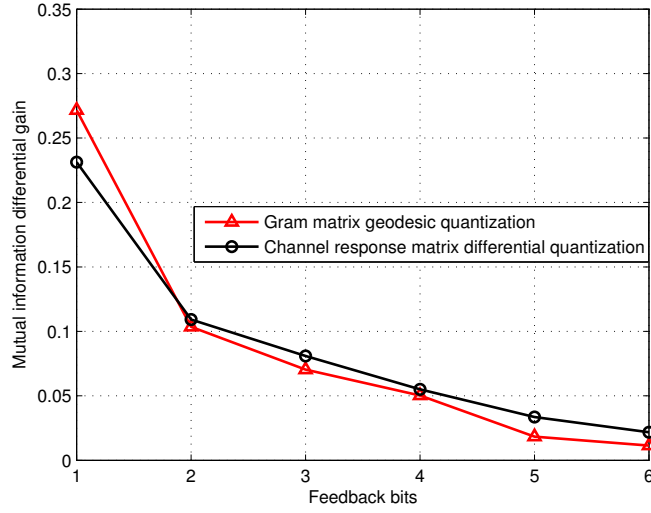


Fig. 3. Differential gain in mutual information in a 3×3 channel with $P_{tx} = 0.5$, $f_D/f_{FB} = 0.05$.

Fig. 5 considers the same scenario, but studies the bit-error-rate (BER) using a BPSK modulation. The curves show a large improvement in the BER when using feedback and in this case the differential scheme also outperforms the

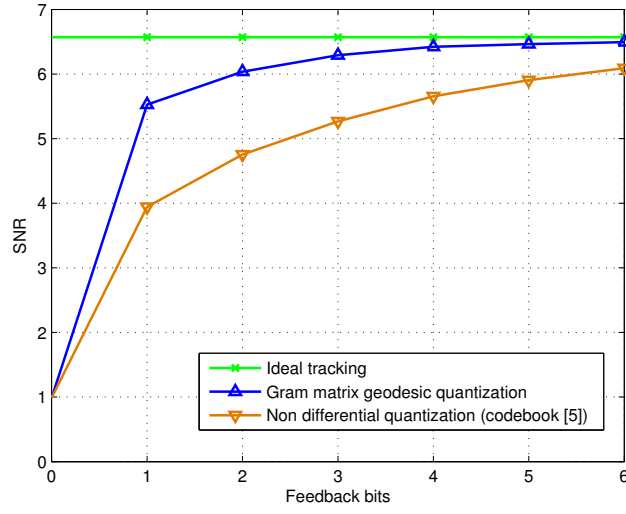


Fig. 4. SNR versus the number of feedback bits.

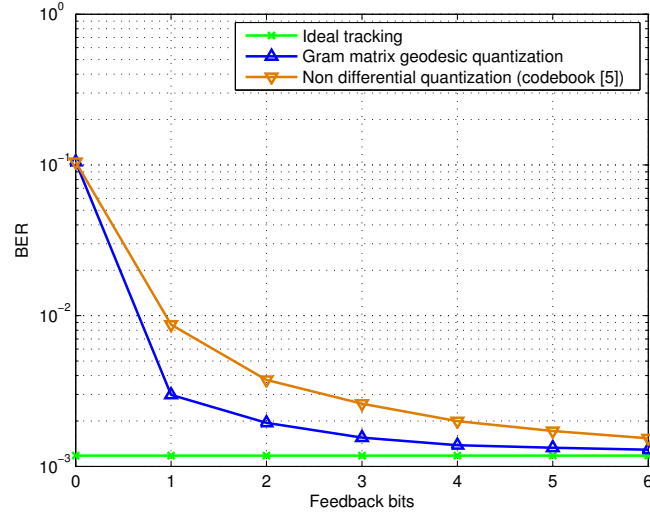


Fig. 5. BER versus the number of feedback bits.

quantization based on codebooks because it exploits the correlation in time of the channel.

5 Resource allocation between data and feedback transmission

In section 4, we have evaluated the benefits of having CSI at the transmitter. In this section we will introduce in the system a variable for the cost of using feedback, in order to study the tradeoff between achievable communication performance and radio resource allocation between the feedback and the data link. In the case of time-division duplex (TDD), the resource that is shared between data and feedback link corresponds to the transmission time, whereas in frequency-division duplexing (FDD) the resource to be shared is the bandwidth. We will now see that to the purpose of resource allocation both schemes are dual, and we will optimize the resource allocation for the general case.

5.1 TDD and FDD systems

In systems where different information streams share the same physical communications link, the available link resources have to be shared. In the case considered in this paper the data and feedback information share the pool of radio resources. We will consider two duplexing schemes: dividing the time axis in different time slots and assigning each slot to the transmission of either data or feedback information (TDD), and dividing the frequency axis in different bandwidth slots, corresponding to feedback or data transmission (FDD). For the equations describing these schemes we use the following notation:

- W_t : total available bandwidth.
- W_d : bandwidth dedicated to transmission of data.
- $W_f = W_t - W_d$: bandwidth dedicated to transmission of feedback.
- T_t : total duration of a time frame.
- T_d : time dedicated to transmission of data.
- $T_f = T_t - T_d$: time dedicated to transmission of feedback.
- E_t : total available energy for the transmission of data.
- N_0 : noise power spectral density (AWGN).
- R_d : rate at which data can be transmitted.
- $\mathbf{H} \in \mathbb{C}^{n_R \times n_T}$: flat fading MIMO channel with n_T and n_R transmit and receive antennas, respectively
- $\mathbf{B}(n) \in \mathbb{C}^{n_T \times n_S}$: linear transmitter matrix that satisfies $\|\mathbf{B}\|_F^2 \leq 1$.

Frequency-division duplex: The FDD scheme features continuous transmission of data and feedback simultaneously, by dividing the total bandwidth available W_t between the data and the feedback link, as depicted in Fig. 6

In such a system, the maximum achievable data rate is given by the following expression, which is an increasing function of the available bandwidth W_d :

$$R_d = W_d \log_2 \det \left(\mathbf{I} + \frac{E_t}{W_d N_0} \mathbf{H}^H \mathbf{H} \mathbf{B} \mathbf{B}^H \right) \quad (\text{bits/s}). \quad (4)$$

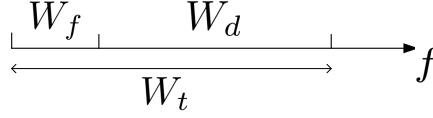


Fig. 6. FDD system model.

Time-division duplex: On the other hand, the TDD scheme makes use of the complete bandwidth to transmit either data or feedback information⁵. The scheduling is performed in the time domain, i.e., there are time slots where all the bandwidth is devoted to sending data and in the other time slots all the bandwidth is dedicated to the feedback link, as depicted in Fig. 7

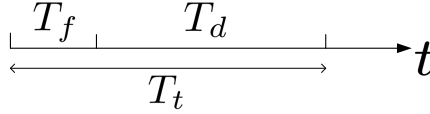


Fig. 7. TDD system model.

In a TDD system, the maximum achievable data rate is an increasing function of the time devoted to transmitting data, and is given by the following expression:

$$R_d = \frac{T_d}{T_t} W_t \log_2 \det \left(\mathbf{I} + \frac{\frac{E_t}{T_d}}{W_t N_0} \mathbf{H}^H \mathbf{H} \mathbf{B} \mathbf{B}^H \right) \quad (\text{bits/s}) \quad (5)$$

General expression (TDD + FDD): As observed in (4) and (5), the expressions of the data rate for both TDD and FDD are dual, and they behave exactly the same as a function of variables T_d and W_d , respectively. It is possible to jointly formulate this dependance (based on (4) and (5)):

$$R_d = \frac{T_d}{T_t} \frac{W_d}{W_t} W_t \log_2 \det \left(\mathbf{I} + \frac{E_t}{T_t W_t N_0} \frac{1}{\frac{T_d}{T_t} \frac{W_d}{W_t}} \mathbf{H}^H \mathbf{H} \mathbf{B} \mathbf{B}^H \right) \quad (\text{bits/s}) \quad (6)$$

The case where $T_d = T_t$ corresponds to FDD, and $W_d = W_t$ corresponds to TDD.

Expression (6), normalized to the bandwidth, can also be written as:

⁵ In the literature it is usually assumed that in TDD systems there is channel reciprocity and therefore feedback is not required. In practical systems, however, the radio frequency (RF) chains have a different response for transmission and for reception. There are two solutions to this issue: one option is to do feedback of the CSI (which includes obviously the effect of the RF chain) and the other option is to perform a calibration of the RF chains for transmission and for reception. In this paper only the feedback solution is considered.

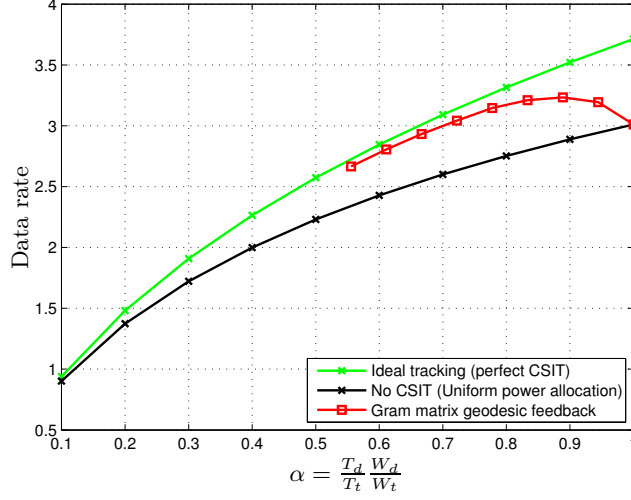


Fig. 8. Data rate versus time dedicated to data and feedback transmission.

$$R_d = \alpha \log_2 \det \left(\mathbf{I} + \frac{snr}{\alpha} \mathbf{H}^H \mathbf{H} \mathbf{B} \mathbf{B}^H \right) \quad (\text{bits/s}), \quad (7)$$

where $\alpha = \frac{T_d W_d}{T_t W_t}$ ($0 \leq \alpha \leq 1$) and $snr = \frac{E_t}{T_t W_t N_0}$.

Following (7), the optimum allocation of resources for data and feedback transmissions can be performed. Fig. 8 shows the results for the case of a 3×3 channel and a normalized doppler frequency of $f_D/f_B = 0.05$. In this simulation it was considered that each bit of feedback used required 1/18 of the available resources in order to obtain an error-free feedback transmission. These results show that, for this particular setup, the largest data rate is achieved by using 2 bits of feedback, i.e. assigning 2/18 of the resources to the feedback link and the remaining 16/18 to the data link.

5.2 A practical case

In this section we study a practical situation of single beamforming, featuring the transmission of frames with a fixed length of 18 bits. 2 of these bits are reserved for control and other information. Of the remaining $L = 16$ bits, N of them are used for feedback and the rest $(L - N)$ are used for data transmission, as depicted in Fig. 9.

We now consider the optimum allocation of bits between the feedback and data links. Note that the transmission of feedback is done through a noise-free and delay-free link, and the transmission of data is done using a QPSK modulation in AWGN.

The data is transmitted in packets of L_p bits. We assume that if there is a transmission error in one of the bits of the packet, the packet is discarded. This

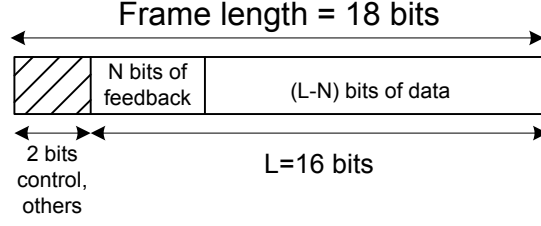


Fig. 9. Frame structure.

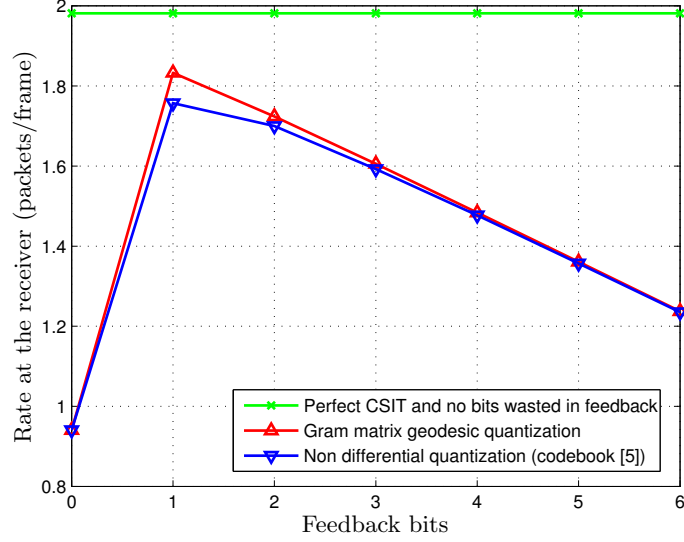


Fig. 10. Average number of data packets received correctly for each frame.

means that on average for each frame the number of packets received without errors is:

$$\frac{L - N}{L_p} (1 - PER) = \frac{L - N}{L_p} (1 - BER)^{L_p}. \quad (8)$$

Following this expression, Fig. 10 shows the packet rate per frame for the case of $P_{tx} = 1$, $L = 16$, $L_p = 8$ and a 3×3 MIMO time variant channel with normalized Doppler of $f_D/f_{FB} = 0.05$. The results are averaged over 1000 channel realizations. The resulting curves show that, for the scenario considered, the best packet transmission rate is achieved when using a feedback load of 1 bit. This holds for the differential and also the non-differential feedback schemes.

6 Conclusions

This work presented an analysis of the resource allocation (time and bandwidth) between the data and the feedback link of a MIMO communication system. It is shown that, since resources for the feedback transmission come at a cost of resources for the data transmission, there is an optimum resource allocation strategy that maximizes system throughput.

We considered non-differential and differential feedback algorithms, with special focus on the later, and performed simulations for different scenarios. The results show that a low rate feedback link is usually enough to provide almost all CSI to the transmitter, and the additional accuracy obtained by increasing the feedback load does not compensate for the loss in bits that would otherwise be used to transmit data.

References

1. Telatar, I.E.: Capacity of multi-antenna Gaussian channels. *European Trans. on Telecommunications*. vol. 10, no. 6, pp. 585–595 (1999)
2. Alamouti, S.M.: A simple transmit diversity technique for wireless communications. *IEEE Journal on Selected Areas in Communications*. vol. 16, no. 8, pp. 1451–1458 (1998)
3. Payaró, M., Palomar, P.: On optimal precoding in linear vector Gaussian channels with arbitrary input distribution. In: *IEEE International Symposium on Information Theory (ISIT'09)*, Seoul (2009)
4. Palomar, D.P., Cioffi, J.M., Lagunas, M.A.: Joint Tx-Rx beamforming design for multicarrier MIMO channels: a unified framework for convex optimization. *IEEE Transactions on Signal Processing*. vol. 51, no. 9, pp. 2381–2401 (2003)
5. Roh, J.C., Rao, B.D.: Design and analysis of MIMO spatial multiplexing systems with quantized feedback. *IEEE Transactions on Signal Processing*. vol. 54, no. 8, pp. 2874–2886 (2006)
6. Sacristán-Murga, D., Pascual-Iserte, A.: Differential feedback of MIMO channel correlation matrices based on geodesic curves. In: *34th IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP'09)*, Taipei (2009)
7. Love, D.J., Heath, R.W., Strohmer, T.: Grassmannian beamforming for multiple-input multiple-output wireless systems. *IEEE Transactions on Information Theory*. vol. 49, no. 10, pp. 2735–2747 (2003)
8. Talih, M.: Geodesic Markov chains on covariance matrices. *Statistical and Applied Mathematical Sciences Institute*. pp. 1–27 (2007)
9. Steele, R., Hanzo, L.: *Mobile Radio Communications*. John Wiley & Sons (1999)
10. Yang, J., Williams, D.B.: Transmission subspace tracking for MIMO systems with low-rate feedback. *IEEE Transactions on Communications*. vol. 55, no. 8, pp. 1629–1639 (2007)
11. Inoue, T., Heath, R.W.: Geodesic prediction for limited feedback multiuser MIMO systems in temporally correlated channels. In: *4th IEEE Radio and Wireless Symposium*, San Diego (2009)
12. Roh, J.C., Rao, B.D.: Efficient feedback methods for MIMO channels based on parameterization. *IEEE Transactions on Wireless Communications*. vol. 6, no. 1, pp. 282–292 (2007)